

Real-time hydrograph modelling in the upper Nysa Kłodzka river basin (SW Poland): a two-model hydrologic ensemble prediction approach

Tomasz Niedzielski¹ · Bartłomiej Miziński¹

Published online: 28 April 2016

© The Author(s) 2016. This article is published with open access at Springerlink.com

Abstract A new system for computing the multimodel ensemble predictions of water levels, known as HydroProg, is applied in the process of real-time modelling and forecasting riverflow in the upper Nysa Kłodzka basin (SW Poland). The HydroProg system automatically produces early warnings against high flows on a basis of prognoses from external hydrologic models which are run in a cloud-like fashion using the real-time hydrometeorological observations. The predictions offered by the models serve as ensemble members and become inputs to HydroProg which does data pre- and post-processing as well as computes multimodel ensemble predictions. The HydroProg system is used in this paper to verify the operational hypothesis that a two-model ensemble hydrologic prediction, solely based on two prognoses computed from the multi- and univariate autoregressive statistical models, reveals better skills than the individual members. The analysis is conducted for all lead times and selected phases of high flow development, including a rising limb of hydrograph but excluding a peak flow. The investigation is carried out using the first HydroProg prototype that works experimentally for the upper Nysa Kłodzka basin (SW Poland). The implementation offers 3-h predictions of water levels, based on the regularly (15 mins) calibrated and updated individual models (the emphasis is put on two models, but other models are also presented for reference) which become inputs to the multimodel ensemble solution (mainly the two-model approach, with the six-model

solution presented for comparison). The most significant high flows that occurred in the study area between 01/09/2013 and 31/08/2015 have been investigated. It has been found that for slightly rising water level the two-model ensemble is recommended which, for a subsequently moderately or rapidly rising limb of hydrograph, should be replaced by the vector autoregressive model.

Keywords Hydrology · Real-time prediction · Multimodel ensemble · Hydrograph · High flow · Poland

1 Introduction

Recent progress in hydrologic prediction is due to hydrological ensemble prediction systems (HEPS) which enable the determination of a hydrograph prognosis on a basis of numerous forecasts. Associated with this is an ensemble “spaghetti” hydrograph (Cloke and Pappenberger 2009; their Fig. 1) that includes individual forecasts, known as ensemble members, the number of which varies across the operational systems and experiments. According to Cloke et al. (2013), the HEPS-based forecasts can be calculated on a basis of numerous approaches, including the most common one, i.e. the use of numerical weather prediction (NWP) based on ensemble prediction systems (EPS) (Buizza et al. 2005). Within such a framework, the EPS-based ensemble members, corresponding to multiple weather forecasts computed under the assumption of various initial conditions of the NWP model, become inputs to a hydrologic model, leading to the determination of numerous hydrologic ensemble members to predict riverflow. The above approach is utilized to support the operational flood forecasting systems (Cloke and Pappenberger

✉ Tomasz Niedzielski
tomasz.niedzielski@uwr.edu.pl

¹ Department of Geoinformatics and Cartography, Faculty of Earth Science and Environmental Management, University of Wrocław, pl. Uniwersytecki 1, 50-137 Wrocław, Poland

Fig. 1 Concept of producing the relationship matrix **R** defined in Eq. 2

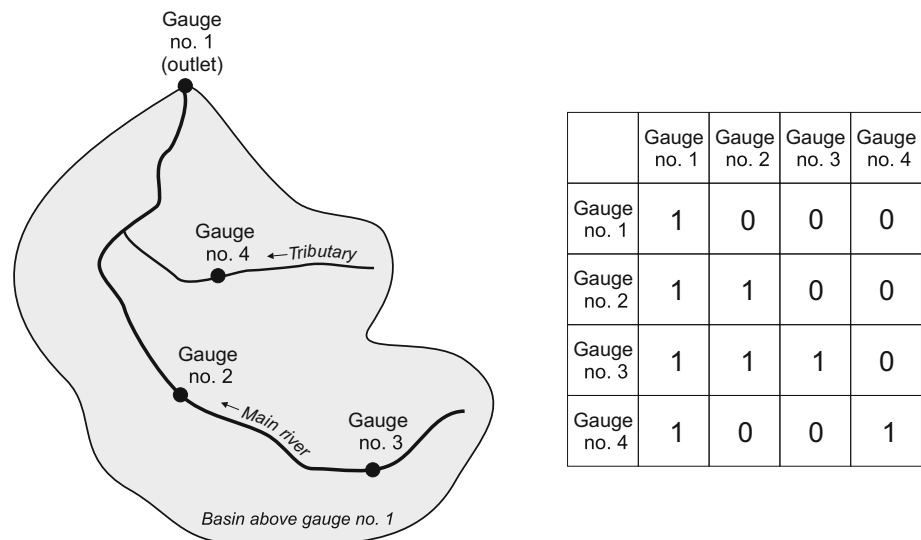


Table 1 List of gauges for which the HydroProg-Kłodzko prototype provides real-time predictions of water level, along with basic information on: the river names, areas of contributing basins and mean elevations of contributing basins

| Gauge | River | Area (km ²) | Mean elevation ^a (m a.s.l.) |
|-------------------|---------------------|-------------------------|--|
| Międzylesie | Nysa Kłodzka | 49.4 | 595.3 |
| Bystrzyca Kłodzka | Nysa Kłodzka | 260.0 | 576.5 |
| Krosnowice | Nysa Kłodzka | 484.2 | 557.0 |
| Kłodzko | Nysa Kłodzka | 1080.0 | 567.7 |
| Bardo | Nysa Kłodzka | 1744.1 | 534.8 |
| Lądek-Zdrój | Biała Łądecka | 162.9 | 765.7 |
| Żelazno | Biała Łądecka | 303.9 | 648.3 |
| Szczytna | Bystrzyca Dusznicka | 100.5 | 671.2 |
| Szalejów Dolny | Bystrzyca Dusznicka | 173.3 | 579.9 |
| Ścinawka Górna | Ścinawka | 398.6 | 523.7 |
| Gorzuchów | Ścinawka | 516.8 | 505.1 |

^a Mean elevation is based on the digital elevation model (DEM), and is computed from all raster cells within the basin

2009; their Table 1) and to enhance the decision support products (Ramos et al. 2007).

Although their considerable potential confirmed in the field of meteorology, the HEPS-based hydrologic prognoses remain new solutions. In order to demonstrate the advantages of HEPS, in 2004 the hydrologic ensemble prediction experiment (HEPEX) has been inaugurated (Franz et al. 2005), and is continuously providing scientists and practitioners with new developments in the field of HEPS (Schaaake et al. 2006; Thielen et al. 2008; Schaaake et al. 2010; Andel et al. 2013). The updated information on HEPS is provided on the HEPEX website www.hepex.org.

There are numerous examples of HEPS. For instance, the mesoscale alpine programme demonstration of probabilistic hydrological and atmospheric simulation of flood events in the Alpine region, known as MAP D-PHASE, has been implemented for the Alps (Zappa et al. 2008). Yet

another example is the HEPS in Sweden where meteorological ensemble forecasts computed by the European Centre for Medium-range Weather Forecasts (ECMWF) become inputs to the hydrologic model known as “Hydrologiska Byråns Vattenbalansavdelning” (HBV) (Olsson and Lindström 2008). In the regional scale, there exist the EXperimental Ensemble Forecast Service (XEFS) and its operational Hydrologic Ensemble Forecast Service (HEFS), both of the US National Oceanic and Atmospheric Administration (NOAA) (Demargne et al. 2013). Recently, Regonda et al. (2013) proposed a new procedure of producing the ensemble hydrologic prediction, the name of which is the hydrologic model output statistics (HMOS). It computes ensemble prognoses on a basis of probability distribution of future riverflow, prediction of rainfall and the up-to-date riverflow data. It is worth noting that the HEPS-based forecasts are useful not only for short- and

medium-term riverflow predictions, but are also valuable in a process of anticipating seasonal extremes (Najafi and Moradkhani 2015). Noteworthy is also the fact that the HEPS concept can go beyond streamflow forecasting and may be utilized for instance to predict a recovery from drought events (DeChant and Moradkhani 2015) and to support snow water equivalent prognoses (Dechant and Moradkhani 2011a).

Less common approach to produce the HEPS-based forecasts is referred to as “multimodelling”. Its concept is based on running numerous hydrologic models which are conceptually unrelated and often independent. The multimodel hydrologic ensemble prognosis is calculated using a weighting strategy that assigns weights to forecasts computed from individual models (e.g. See and Abrahart 2001; Ajami et al. 2006; Yan and Moradkhani 2016). The multimodel hydrologic ensemble approach has been shown to produce more accurate riverflow predictions than forecasts based on single hydrologic models fed by meteorological weather ensemble prognoses (Duan et al. 2007; Velázquez et al. 2011). It is also known, however, that skills of multimodel hydrologic predictions differ along with numerous parameters, such as for instance lead time and input data pre-processing. Thus, HEPS solutions do not always offer a pronounced improvement of performance of individual multimodel ensemble members (Bohn et al. 2010). There is no clear picture whether HEPS should be fed by all members or just by their specific subsets (Seiller et al. 2012). To compute the most skillful HEPS-based forecast it is recommended to: select the appropriate ensemble calculation method and choose the appropriate ensemble members from a group of available models (Kumar et al. 2015). In addition, the performance HEPS depends on basin areas and its geographical setting, as shown by Alfieri et al. (2014) for the European flood awareness system (EFAS). Moreover, initial conditions influence the performance of hydrologic ensemble predictions (Wood and Lettenmaier 2008; Li et al. 2009; DeChant and Moradkhani 2011b).

Recently, a new experimental real-time HEPS, based on the concept of multimodelling, has been prototyped. The system, known as HydroProg, serves as a real-time early warning platform that integrates observational networks with hydrologic models run externally by various modellers/teams, combining the individual hydrologic ensemble members into the multimodel hydrologic ensemble forecast (Niedzielski et al. 2014). The concept of keeping models outside of the forecasting systems has been used in the Delft-FEWS solution (Werner et al. 2013). The HydroProg system has recently been implemented by the authors of this paper. The system has been deployed in the Nysa Kłodzka basin (SW Poland), and this implementation is hereinafter referred to as HydroProg-Kłodzko. In SW

Poland, where rivers are highly flood-prone (Dubicki et al. 2005), there have been numerous attempts to produce riverflow predictions (e.g. Butts et al. 2007), including running independent models in order to compare their predictive performance (Butts et al. 2006). Since Central Europe is particularly vulnerable to floods, several attempts have recently been made to employ HEPS for flash flood early warning (Raynaud et al. 2015). Although there were numerous modelling exercises of this type carried out in SW Poland, the only real-time system that enables a comparison between accuracies offered by various hydrologic models is HydroProg-Kłodzko.

The objective of this paper is to show the performance of the real-time HydroProg-based multimodel hydrologic ensemble predictions of water level in relation to skills of prognoses based on individual ensemble members. Rapid (updated every 15 mins), short-term (lead times up to 3 h) water level forecasts produced by the HydroProg-Kłodzko prototype are examined. Since empirical models are still often used in real-time forecasting system (Leedal et al. 2013; Smith et al. 2014), the investigation is limited to two data-based methods (Niedzielski 2007, 2010): the univariate autoregressive (AR) and multivariate vector autoregressive (VAR) models. The operational hypothesis to be verified in this paper states that the two-model ensemble prognosis of riverflow is more accurate than the two individual predictions computed by the two simple data-based hydrologic models, and it is also checked if the latter is the case for all lead times as well as for phases of high flow development, including a rising limb of hydrograph but excluding a peak flow. In addition, the paper aims to experimentally verify if the simple two-model approach performs better than several individual data- and physically-based models and their multimodel ensembles. Thus, given the two-model proposition and simplicity the paper discusses if more and less complicated solutions perform better.

2 Methods

2.1 The HydroProg-Kłodzko prototype

In December 2011, activities towards a novel HEPS commenced at the University of Wrocław, Poland, and their driving force was the project entitled “System supporting a comparison of hydrologic predictions”, funded by the National Science Centre, Poland. The system itself, abbreviated with the acronym HydroProg, is a general infrastructure that provides rapid (15-min calibration and update) and short-term (lead times up to 3 h, with the 15-min intermediate step) integration between hydrometeorological observational networks (hydrologic and

weather gauges) and hydrologic models equipped with predictive skills. The system acts as a real-time early warning infrastructure, with its most fundamental element which is the real-time multimodel hydrologic ensemble generator (Niedzielski et al. 2014). The system: (1) collects hydrometeorological data in real time from data provider, (2) very quickly performs simple data pre-processing (transferring data to UTC time to avoid the problem of setting the clocks back and forward according to the daylight saving time rules, approximating uneven times of observations to get equal time resolution, removing outliers due to sensor failure or other unexpected situations), (3) immediately serves the pre-processed data to modelers based at different groups or institutions (a cloud-like concept), (4) collects predictions computed by modelers and estimates their errors to calculate weights associated with every single ensemble member, (5) generates the multimodel ensemble prognosis in real time using the above-mentioned weights which are updated in real time.

Although the HydroProg system forms a general prediction engine, its prototype has been deployed for the Nysa Kłodzka basin (SW Poland). The HydroProg-Kłodzko prototype comprises: the HydroProg engine (integration of models and computation of multimodel hydrologic ensemble forecasts) and the real-time web map service (visualization of prognoses on user demand). The latter is available at www.klodzko.hydroprog.uni.wroc.pl.

2.2 Hydrologic models

For the purpose of this paper, two data-based hydrologic models have been selected, and they are ensemble members of the HydroProg-Kłodzko implementation since launch of the system in August 2013. Only two models are investigated because they are the only ensemble members which are run uninterruptedly over 2 years of the experiment. Thus, a large database of prediction errors is available, and this allows one to infer forecast skills of the two models, with an emphasis put on how ensemble weights react to high flow conditions.

2.2.1 Multivariate (vector) autoregressive model

Let us consider a multivariate riverflow series $\mathbf{x} = (\mathbf{x}_1, \dots, \mathbf{x}_n)$ of length n , composed of the following vectors:

$$\mathbf{x}_t = \begin{bmatrix} x_t(1) \\ \vdots \\ x_t(m) \end{bmatrix}, \quad (1)$$

where $x_t(j)$, $t = 1, \dots, n$ and $j = 1, \dots, m$, corresponds to river stage measured at t -th time step (in the HydroProg-Kłodzko prototype the time step is equal to 15 mins) at j -th gauge.

In order to account for multivariate dependencies which are due to the river network setting, the following relationship matrix $\mathbf{R} = \{r_{ij}\}_{1 \leq i, j \leq m}$ is used:

$$\mathbf{R} = \begin{bmatrix} r_{11} & \dots & r_{1m} \\ \vdots & \ddots & \vdots \\ r_{m1} & \dots & r_{mm} \end{bmatrix}, \quad (2)$$

where $r_{ij} = 0$ when water flowing through i -th gauge never reaches j -th gauge, and $r_{ij} = 1$ when water flowing through i -th gauge reaches j -th gauge (Fig. 1), and this occurs with certain delay. This approach constrains the dimension of the model, since for a given j -th outlet it is necessary to consider solely spatially-interrelated signals at upstream gauges numbered as i for which $r_{ij} = 1$.

Hence, for a given gauge j , the actual number of inter-related gauges is usually lower than m . Let us assume that for a given j -th gauge, the condition $r_{ij} = 1$ is fulfilled $m(j)$ times, for $i = 1, \dots, m$, and $m(j) \leq m$. Then for each gauge j , the following time series $\mathbf{x}^{(j)} = (\mathbf{x}_1^{(j)}, \dots, \mathbf{x}_n^{(j)})$ is considered, for which

$$\mathbf{x}_t^{(j)} = \begin{bmatrix} x_t^{(j)}(1) \\ \vdots \\ x_t^{(j)}(m(j)) \end{bmatrix}, \quad (3)$$

where for each $k = 1, \dots, m(j)$:

$$x_t^{(j)}(k) = x_t(i) \quad \text{if} \quad \left(r_{ij} = 1 \quad \text{and} \right. \\ \left. k = \#\{g : g \leq i \wedge r_{gj} = 1\} \right). \quad (4)$$

According to Niedzielski (2007), in order to remove linear trends and produce a residual series $\mathbf{y}^{(j)} = (\mathbf{y}_1^{(j)}, \dots, \mathbf{y}_{n-1}^{(j)})$ the use is made of the 1-lag difference operator given by:

$$\mathbf{y}_t^{(j)} = \begin{bmatrix} x_{t+1}^{(j)}(1) - x_t^{(j)}(1) \\ \vdots \\ x_{t+1}^{(j)}(m(j)) - x_t^{(j)}(m(j)) \end{bmatrix}, \quad (5)$$

for $t = 1, \dots, n-1$. Note that \mathbf{x} , $\mathbf{x}^{(j)}$ and $\mathbf{y}^{(j)}$ are trajectories of vector time series models \mathbf{X} , $\mathbf{X}^{(j)}$ and $\mathbf{Y}^{(j)}$, respectively.

Let us assume that the differenced data $\mathbf{y}^{(j)}$ has zero-mean for each gauge (for fixed row). The stationary time series $\mathbf{Y}^{(j)}$ is a zero-mean vector autoregressive process of order p , known also as VAR(p), if it satisfies the following equation:

$$\mathbf{Y}_t^{(j)} = A_1 \mathbf{Y}_{t-1}^{(j)} + \dots + A_p \mathbf{Y}_{t-p}^{(j)} + \mathbf{E}, \quad (6)$$

where $\mathbf{Y}_t^{(j)}$ is a random vector that corresponds to $\mathbf{Y}^{(j)}$ at time step t ; A_d , for $d = 1, \dots, p$, is a matrix composed of autoregressive coefficients; \mathbf{E} is a white noise vector with

mean 0 and covariance matrix \mathbf{C} . There are several statistics to estimate p , but in this paper the use is made of the Akaike information criterion (AIC) (Akaike 1971), and the selection is done from integers limited by the arbitrarily-chosen maximum reasonable lag. Likewise, coefficient matrices may be estimated using numerous approaches (Lardies 1996; Neumaier and Schneider 2001), but herein they are estimated using the ordinary least squares (OLS) procedure. The predictions with lead times bigger than 1 step are calculated in an iterative way, i.e. 1-step prognosis is attached to observational data and such an extended time series is substituted to prediction equation, and so on.

2.2.2 Univariate autoregressive model

The use is also made of a univariate riverflow series $x = (x_1, \dots, x_n)$ of length n which—for a given gauge j and $t = 1, \dots, n$ —is composed of elements $x_t = x_t(j)$. The residual differenced time series $y = (y_1, \dots, y_{n-1})$ is computed as $y_t = x_{t+1} - x_t$ for $t = 1, \dots, n - 1$. Similarly to the multivariate case, x and y are trajectories of time series models X and Y , respectively.

The second model used in the present study is the univariate autoregressive model which is based on a stationary autoregressive process of order p , $AR(p)$, which in a zero-mean case satisfies the following equation:

$$Y_t = a_1 Y_{t-1} + \dots + a_p Y_{t-p} + Z_t, \quad (7)$$

where Y_t is a stochastic process at time step t ; a_d , for $d = 1, \dots, p$, are autoregressive coefficients; Z_t is a sequence of uncorrelated random variables of mean 0 and variance σ^2 . The order p is chosen following the AIC statistics, and autoregressive coefficients are estimated using the Yule-Walker method (e.g. Brockwell and Davis 1996).

2.3 Hydrologic two-model ensemble predictions

The use of VAR and AR models and the derivation of the two-model ensemble predictions are graphically presented in Fig. 2. The hydrologic ensemble predictions used in the HydroProg-Kłodzko prototype are based on multimodelling. Let us denote the two predictions, i.e. VAR- and AR-based ones with the maximum lead time of 12 steps ($12 \times 15 \text{ min} = 3 \text{ h}$), as $P_{VAR}(t, l)$ and $P_{AR}(t, l)$, respectively, where t is a moment expressed in 15-min steps when forecasts are computed (time of last observation) and $l = 1, \dots, 12$ corresponds to a specific prediction horizon. Then, an ensemble prediction $P_{ENS}(t, l)$ is computed as a convex combination of individual predictions which serve as ensemble members, namely:

$$P_{ENS}(t, l) = w_{VAR}(t, l)P_{VAR}(t, l) + w_{AR}(t, l)P_{AR}(t, l), \quad (8)$$

where $w_{VAR}(t, l)$ and $w_{AR}(t, l)$ are time- and horizon-dependent weights for individual ensemble members, and for fixed t and l the following condition holds $w_{VAR}(t, l) + w_{AR}(t, l) = 1$. The weights are estimated using positive prediction error statistics, such as root mean square error (RMSE) or mean absolute error (MAE), with the following formulae:

$$w_{VAR}(t, l) = \frac{\frac{1}{E_{VAR}(t, l)}}{\frac{1}{E_{VAR}(t, l)} + \frac{1}{E_{AR}(t, l)}}, \quad w_{AR}(t, l) = \frac{\frac{1}{E_{AR}(t, l)}}{\frac{1}{E_{VAR}(t, l)} + \frac{1}{E_{AR}(t, l)}}, \quad (9)$$

where $E_{VAR}(t, l)$ ($E_{AR}(t, l)$) is a positive VAR-based (AR-based) prediction error of l -step forecast computed in the interval $[t - 2879, t]$ (holds from 15-min time step and 30-day moving window). In the HydroProg-Kłodzko prototype the use is made of the RMSE statistics to build weights $E_{VAR}(t, l)$ and $E_{AR}(t, l)$. The choice of the 30-day moving window is arbitrary, but allows for rapid modifications of weights during high flows when certain model starts producing inaccurate forecasts. Hence, the impact of individual members on the multimodel ensemble solution is updated in real time.

3 Geographical and experimental setting

3.1 Study area

The HydroProg-Kłodzko prototype is implemented for the Kłodzko Valley (SW Poland) which is a mid-mountain abasement. The valley itself is surrounded by a few mountain chains, the topographical and geological settings of which differ significantly. More specifically, the HydroProg-Kłodzko covers the Nysa Kłodzka basin above the outlet in Bardo (approximately 260 m a.s.l.). The basin (with area of 1744 km²) covers a larger area than the Kłodzko Valley itself, including not only Polish parts outside the valley but also Czech parts (Fig. 3). The highest elevation within the basin is equal to 1425 m a.s.l. The Nysa Kłodzka river is a main river in the valley, with several key tributaries in the study area: Wilczka, Bystrzyca, Biała Łądecka, Bystrzyca Dusznicka and Ścinawka. The Nysa Kłodzka river is a left tributary of the Odra river, the second largest river in Poland. The main town in the study area is Kłodzko, located centrally in the valley. The town experienced severe floodings (Kasprzak 2010), including recent 1997, 1998 and 2010 events. Not uncommonly, damaging flash floods occur in the study area, and this causes a need for developing real-time rapid warning systems to support authorities and citizens.

Fig. 2 Computation of a two-model hydrologic ensemble prediction—a flowchart

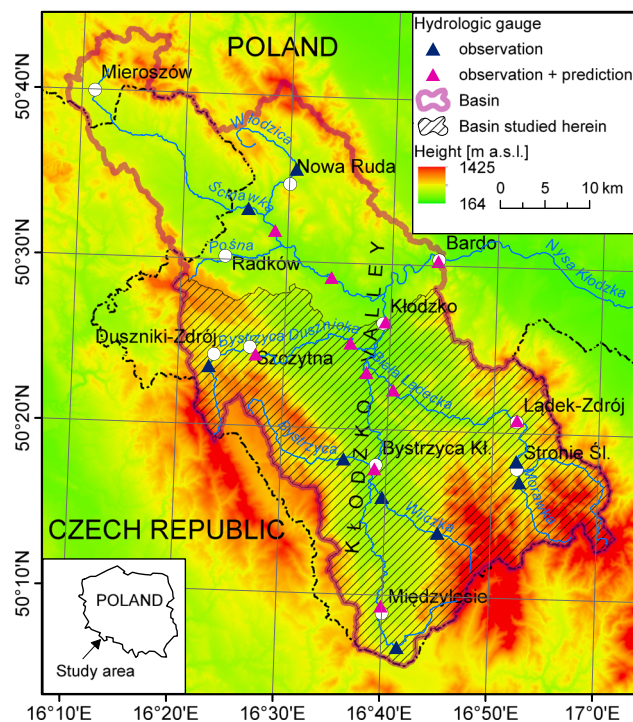
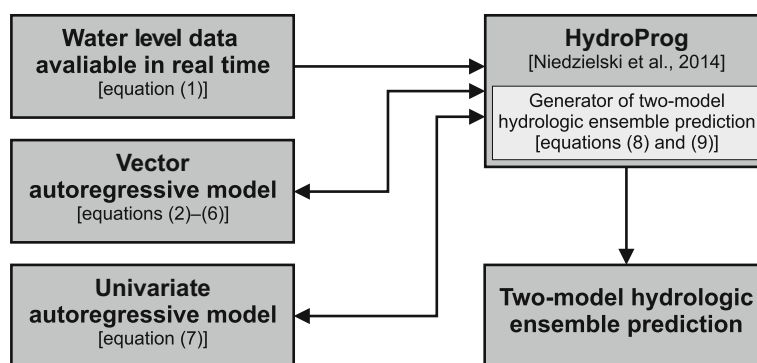


Fig. 3 Map of study area

With the largest Nysa Kłodzka basin above the gauge in Bardo, the HydroProg-Kłodzko prototype uses a set of sub-basins with the associated outlets equipped with automatic gauges along: Nysa Kłodzka (Międzyzłesie, Bystrzyca Kłodzka, Krosnowice, Kłodzko, Bardo), Biała Łądecka (Łądek-Zdrój, Żelazno), Bystrzyca Dusznicka (Szczytna, Szalejów Dolny), Ścinawka (Ścinawka Górna, Gorzuchów). Table 1 juxtaposes the outlets along with the information on their areas and mean elevations.

The rivers themselves form a mixture of regulated and natural channels, with only two major dry detention reservoirs that may control runoff during high flows: in Stronie Śląskie (max area of 24.5 ha), built in the lower part of the stream of Morawka, and in Międzygórze (max area of 6 ha), located in upper part of Wilczka. The maximum discharge, when water is released from the reservoir,

is approximately equal to 37.1 and 10.0 m³/s for Stronie Śląskie and Międzygórze, respectively (Lenar-Matyas et al. 2009). Filling the dry reservoirs occurs rarely, and typically the presence of the two dams does not influence discharge downstream (the river channel is located in the bottom of the dry reservoirs and water flows almost undisturbed). It is believed that the human-controlled dam management does not undermine the hydrologic real-time prediction experiment discussed in this paper since, similarly to Sen and Niedzielski (2010), it has been found that the percentage of maximum water storage in upstream reservoirs in respect to the total runoff is low. In addition, the analysis of daily reports from the two dry detention reservoirs shows that, over the study period, the only usage took place on 14 and 15 September 2013 in the Międzygórze dam. However, during that high flow event the volumes were low in respect to the total capacity, i.e. 11,000 and 19,000 m³ (1.3 and 2.2 % of the total volume), for the 2 days respectively.

For the purpose of the exercise carried out in this paper an emphasis is placed on the Kłodzko basin which belongs to the above-mentioned bigger basin above the outlet in Bardo. The gauge in Kłodzko is selected in this study as the most representative because the site is centrally-located in the Kłodzko Valley, with riverflow influenced by the extensive explanatory information from upstream gauges (Fig. 3).

3.2 Data

Hydrometeorological data are acquired from the local system for flood monitoring (Lokalny System Ochrony Przeciwpowodziowej—LSOP). The LSOP infrastructure comprises: 22 automatic hydrologic gauging stations that observe water level (20 in the studied basin), 16 automatic weather stations (12 in the studied basin) that observe several meteorological variables including precipitation, and the management centre in Kłodzko. The data are being gathered in real time, approximately every 15 mins, however sampling interval varies and may be slightly longer or

shorter than 15 mins. Along with the real-time measurements, historical data collected since 2002 (launch of the LSOP system) have become available for modelling purposes, specifically to run hydrograph models. The HydroProg system includes a few data processing modules, one of which produces so called “processed data”, i.e. equally-sampled (precise 15-min time step is needed for modelling, the UTC time is needed to avoid summer and winter clock changes) and artifact-filtered (minimized number of outliers, interpolated values when they are not available) data. All historical and real-time data are stored in the PostgreSQL¹ database.

Among 20 hydrologic gauges that operationally work in the studied basin, 11 sites for which the HydroProg-Kłodzko prototype continuously calculates water level predictions have been selected. The map presented in Fig. 3 shows the spatial distribution of all gauges, among which there are prediction gauges (Table 1). In addition, Table 2 shows the relationship matrix, $\mathbf{R} = \{r_{ij}\}_{1 \leq i, j \leq 20}$, which is based on relations that can be inferred from Fig. 1. The matrix records the information on network-like relationships between hydrographs in the basin, and such relations are intrinsically associated with the delay of the hydrologic signal due to propagation of the flood wave. This has been highlighted on the real-data example from the high flow on 16–18/05/2014 (Fig. 4), in which the hydrograph from one of the most upstream gauges along the Nysa Kłodzka river (Międzyzylesie) is compared with water level time series from the downstream gauge along the same river (Kłodzko). As mentioned above, the detailed analyses presented in this paper are limited to Kłodzko. Gericke and Smithers (2014) provide a few definitions of a concentration time, and one of them (“the time from the start of total runoff (rising limb of hydrograph) to the peak discharge of total runoff”) has been used in this paper. It has been found that the studied catchments respond rapidly, i.e. from approximately 1 h to 1 day (Table 3).

The HydroProg-Kłodzko prototype utilizes hydrometeorological data that span the time interval from 2005 until now, which corresponds to the large database consisting of over 300,000 time steps, 15-min length each. The adopted VAR and AR methods process only hydrologic data on water level, hence precipitation data are omitted in the case of the two models (however, for a few techniques used within the HydroProg-Kłodzko prototype rain data are

required). Water level time series, before being predicted, are pre-processed in order to reduce a number of outliers present in raw data. For this purpose, the Rosner test for outliers is used (Rosner 1983), the method that is widely accepted in hydrological applications (McCuen 2003).

The performance of each of the VAR-, AR- and ENS-based approaches to predict high flows is assessed. Several high flows have been selected for the purpose of this experiment, and the choice was based on the identification of dates when the 5 % empirical quantiles of water level were exceeded. As a result, the following most significant high flows that occurred between 01/09/2013 and 31/08/2015 (dates of rising limbs of hydrographs are given) have been selected: 14 September 2013, 9 December 2013, 23 March 2014, 16 May 2014, 27 May 2014, 31 July 2014, 1/2 September 2014, 11 September 2014, 22/23 October 2014, 19/20 December 2014, 9 January 2015, 29/30 March 2015. Although there have been six ensemble members available in the HydroProg-Kłodzko prototype, VAR and AR methods remain the only solutions which were available for the preselected high flow episodes (Table 4).

4 Results

The objective of the work is to evaluate the performance of VAR, AR and ENS prediction methods in forecasting hydrograph at Kłodzko gauge in two specific moments: before high flow (before rising limb of hydrograph) and at the beginning of high flow (rising limb of hydrograph), as depicted in Fig. 5. The visual analysis of Figs. 6, 7, 8, 9, 10 and 11 allows one to evaluate the performance of a given prognostic model during high flows.

The VAR- and ENS-based predictions are found to be the most skillful, however their performance depends on the starting prediction times, i.e. on moments when the prognoses have been issued.

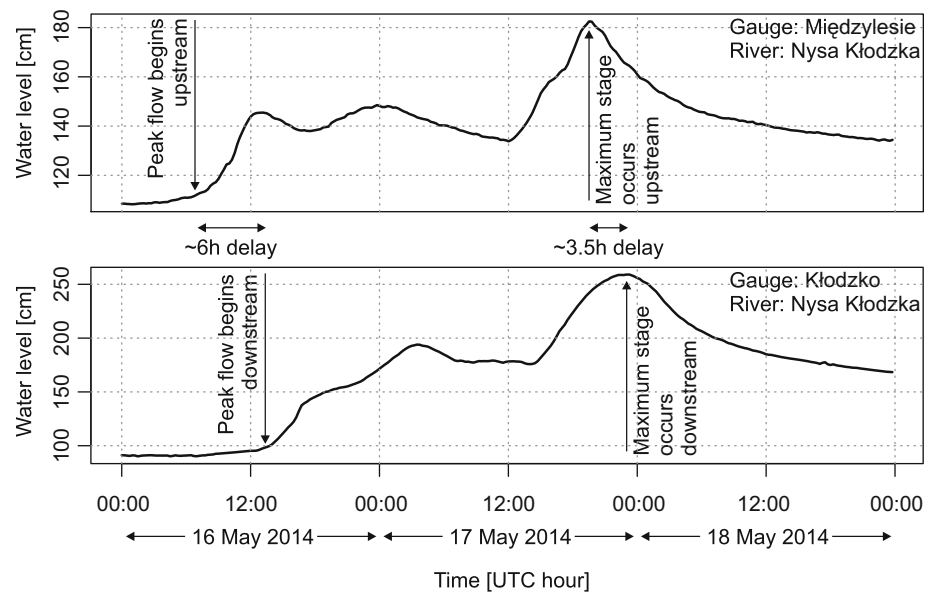
1. When forecasts are issued before a high flow to predict a slightly rising water level in the 3-h horizon with the 15-min temporal resolution, usually the ENS method is very accurate for at least 50 % of lead times (scenario 1).
2. When forecasts are issued at the initial phase of a high flow to predict a rapidly rising limb of hydrograph in the 3-h horizon with the 15-min temporal resolution, usually either
 - (a) the ENS approach is accurate for short lead times (up to 1.5 h into the future) and the VAR method is accurate for longer lead times (scenario 2a) or
 - (b) VAR is superior over both AR and ENS for all lead times (scenario 2b).

¹ PostgreSQL is an open source database, <http://www.postgresql.org/>.

Table 2 Relationship matrix described in Eq. 2, produced for the upper Nysa Kłodzka basin, upstream the outlet in Bardo

| | Bystrzyca Kłodzka | Gorzuchów | Szalejów Dolny | Lądek- Zdrój | Kłodzko | Krosnowice | Żelazno | Międzylesie | Wilkanów | Boboszków |
|---------------------------|----------------------|--------------------|---------------------------|--------------------|----------|------------|-------------------|-------------|--------------|--------------------|
| Bystrzyca Kłodzka | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| Gorzuchów | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Szalejów Dolny | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Lądek-Zdrój | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Kłodzko | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Krosnowice | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| Żelazno | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 |
| Międzylesie | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 |
| Wilkanów | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| Boboszków | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 1 |
| Międzygórze | 1 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 |
| Stronie Śląskie | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Stronie Śląskie zaporą | 0 | 0 | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 0 |
| Duszniki-Zdrój | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Szczytna | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| Tłumaczów | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Ścinawka Górna | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Bardo | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Nowa Ruda | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| Stara Bystrzyca | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 0 | 0 |
| | Międzygórze | Stronie Śląskie | Stronie Śląskie zaporą | Duszniki- Zdrój | Szczytna | Tłumaczów | Ścinawka Górna | Bardo | Nowa Ruda | Stara Bystrzyca |
| Bystrzyca Kłodzka | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Gorzuchów | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Szalejów Dolny | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Lądek-Zdrój | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Kłodzko | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Krosnowice | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Żelazno | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Międzylesie | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Wilkanów | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Boboszków | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Międzygórze | 1 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Stronie Śląskie | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Stronie Śląskie zaporą | 0 | 1 | 1 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Duszniki-Zdrój | 0 | 0 | 0 | 1 | 1 | 0 | 0 | 1 | 0 | 0 |
| Szczytna | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 |
| Tłumaczów | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 | 0 |
| Ścinawka Górna | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 0 | 0 |
| Bardo | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 0 |
| Nowa Ruda | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 1 | 1 | 0 |
| Stara Bystrzyca | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 |

Fig. 4 Delay of the hydrologic signal observed during the development of the high flow on 16–18 May 2014, record between the upstream gauge (Międzyzlesie) and the downstream gauge (Kłodzko)



4.1 Prediction of hydrograph before evident signatures of high flow

The vast majority of studied cases (11/12) fit the scenario 1. For seven high flow episodes, the two-model ensemble prediction approach was found to be the most skillful for at least 75 % of lead times. In the case of four events, the ENS-based forecasts performed best for 50–58 % of lead times. One high flow episode did not follow the scenario 1 (Table 5).

4.1.1 Scenario 1 for at least 75 % of lead times

Spread of the VAR and AR predictions computed at 05:15 UTC on 9 December 2013 became a virtue as the two-model ensemble almost perfectly fitted the hydrograph in its entire 3-h interval 05:15–08:15 UTC (Fig. 6b). Low weights for VAR predictions at long lead times (approximately 0.3 for 3 h into the future) successfully eliminated the impact of VAR overestimation on the two-model ensemble (Fig. 6a). Similarly, the VAR- and AR-based predictions issued at 09:15 UTC on 16 May 2014 revealed a big spread around the hydrograph which was approximately equal to 7 and 3 cm, respectively (Fig. 6e). The weights preferred the univariate AR solution, and the VAR-associated overestimating prognoses received weights of approximately 0.3 (Fig. 6d) which led to production of a very accurate ENS prognosis of water level for the vast majority of lead times.

The VAR- and AR-based prognoses computed at 12:45 UTC on 27 May 2014 were approximately equally weighted before the high flow, with slight dominance of

AR for lead times of 15 mins and tiny dominance of VAR for lead times of 3 h (Fig. 7b). When the predictions were weighted using almost equal weights (Fig. 7a), the two-model correctly anticipated the hydrograph for the vast majority of lead times. Similarly, such a weighting eliminated the spread of prognoses computed using the two methods at 13:00 UTC on 31 July 2014 (Fig. 7d), and the resulting two-model ensemble perfectly fitted the hydrograph for all lead times (Fig. 7e).

Although the two-model ensemble was found to produce the most accurate prognosis of slightly rising hydrograph in its entire 3-h interval from 21:15 UTC on 1 September 2014 onwards (Fig. 8b), the weighting worked better for short lead times than for long lead times. Indeed, the AR-based 3-h prediction received higher rank than the prognosis computed using the VAR method (Fig. 8a) which—along with the fact that magnitude of VAR overestimation of the true hydrograph was smaller than the corresponding magnitude for the AR forecast—led to certain mismatch between the ENS-based prognosis for long lead times and observed data. An opposite situation regarding weights occurred for the prognoses of slow rise of water level computed at 09:45 UTC on 11 September 2014. The VAR approach was slightly favoured, particularly for long lead times (Fig. 8d). The VAR-based prognosis performed better (but still overestimated the hydrograph) than the AR solution (which still underestimated the hydrograph) and—given the above-mentioned weights—the resulting two-model ensemble offered a very accurate prediction of water level for the vast majority of lead times (Fig. 8e).

The magnitudes of mismatches were dissimilar for the two solutions computed at 20:45 UTC on 29 March 2015,

Table 3 Concentration time computed for 11 gauges and 12 high flows

| Gauge | Concentration time $T_C(d)^a$ (hours) | | | | | | |
|-------------------|---------------------------------------|----------------------|----------------------|----------------------|----------------------|----------------------|----------------------|
| | 14/09 2013 | 09/12 2013 | 23–24/03 2014 | 16–17/05 2014 (A) | 16–17/05 2014 (B) | 27–28/05 2014 (A) | 27–28/05 2014 (B) |
| Międzyzlesie | 13.50 | 12.50 | 12.75 | 5.25 | 7.25 | 1.25 | 2.50 |
| Bystrzyca Kłodzka | 10.75 | 25.50 | 8.75 | 14.50 | 7.50 | 1.25 | 2.25 |
| Krosnowice | 9.25 | – | – | 13.25 | 8.00 | 4.50 | 2.50 |
| Kłodzko | 10.00 | 26.25 | 15.00 | 14.75 | 8.25 | 5.50 | 9.00 |
| Bardo | – | 19.25 | 13.50 | 13.00 | 11.50 | 4.00 | 7.25 |
| Lądek-Zdrój | 11.50 | 19.00 | 11.50 | 18.25 | 7.25 | ← 10.25 → | |
| Żelazno | 10.50 | 20.25 | 12.75 | 15.50 | 6.75 | 4.00 | 10.00 |
| Szczytna | 11.00 | 23.00 | 13.25 | ← 8.75 → | | 2.50 | 6.00 |
| Szalejów Dolny | 7.25 | 22.75 | 13.00 | ← 7.00 → | | 4.75 | 5.50 |
| Ścinawka Górna | 12.75 | 17.75 | 14.25 | 8.75 | 6.75 | ← 4.75 → | |
| Gorzuchów | 14.75 | – | – | 14.25 | 5.75 | ← 6.50 → | |
| | 31/07 2014 | 01–02/09 2014 | 11–12/09 2014 | 22–23/10 2014 | 19–21/12 2014 (A) | 19–21/12 2014 (B) | 19–21/12 2014 (C) |
| Międzyzlesie | 5.75 | 13.25 | 24.00 | 8.50 | 10.25 | 6.50 | 7.25 |
| Bystrzyca Kłodzka | 4.25 | 13.75 | 26.00 | 9.75 | 11.25 | 6.25 | 7.50 |
| Krosnowice | 5.50 | 15.25 | 14.00 | 7.50 | 10.50 | 5.50 | 7.75 |
| Kłodzko | 6.00 | 18.25 | 24.75 | 7.75 | ← 24.25 6.50 → | | |
| Bardo | – | 13.50 | 21.75 | 11.00 | ← 19.75 5.75 → | | |
| Lądek-Zdrój | 6.25 | 6.75 | 14.50 | 10.25 | ← 11.25 → | | |
| Żelazno | 7.75 | 13.25 | 20.75 | 8.25 | ← 15.50 → | | |
| Szczytna | 12.00 | 8.50 | 16.50 | 7.00 | ← 24.50 → | | |
| Szalejów Dolny | 9.00 | 10.50 | 19.75 | 9.75 | ← 27.00 → | | |
| Ścinawka Górna | 2.50 | 19.50 | 10.25 | 9.00 | ← 20.75 8.50 → | | |
| Gorzuchów | 3.50 | 11.25 | 24.50 | 8.00 | – | – | – |
| | 08–09/01 2015 (A) | 08–09/01 2015 (B) | 08–09/01 2015 (C) | 29–31/03 2015 (A) | 29–31/03 2015 (B) | 29–31/03 2015 (C) | |
| Międzyzlesie | 5.50 | 3.75 | 6.50 | 12.75 | 2.75 | 12.75 | |
| Bystrzyca Kłodzka | 5.75 | 5.75 | 6.00 | 9.25 | 3.75 | 8.50 | |
| Krosnowice | 6.50 | 6.50 | 5.00 | 7.75 | 4.75 | 11.25 | |
| Kłodzko | 9.75 | 6.25 | 6.00 | 8.50 | 4.25 | 11.75 | |
| Bardo | 7.25 | 5.75 | 11.25 | ← 7.25 8.00 → | | | |
| Lądek-Zdrój | 6.00 | 5.25 | 14.00 | ← 13.25 11.50 → | | | |
| Żelazno | ← 9.25 21.00 → | | | ← 20.50 12.25 → | | | |
| Szczytna | ← 7.75 22.75 → | | | ← 12.00 10.50 → | | | |
| Szalejów Dolny | 6.50 | 5.00 | 11.00 | ← 10.75 11.00 → | | | |
| Ścinawka Górna | 8.00 | 6.75 | 12.25 | ← 12.00 15.00 → | | | |
| Gorzuchów | 5.25 | 5.50 | 7.25 | ← 9.75 10.75 → | | | |

^a Time from beginning of total runoff to peak discharge of total runoff (Gericke and Smithers 2014). (A), (B), (C)—Symbols of two or three evident local peak flows present within a long high flow episode

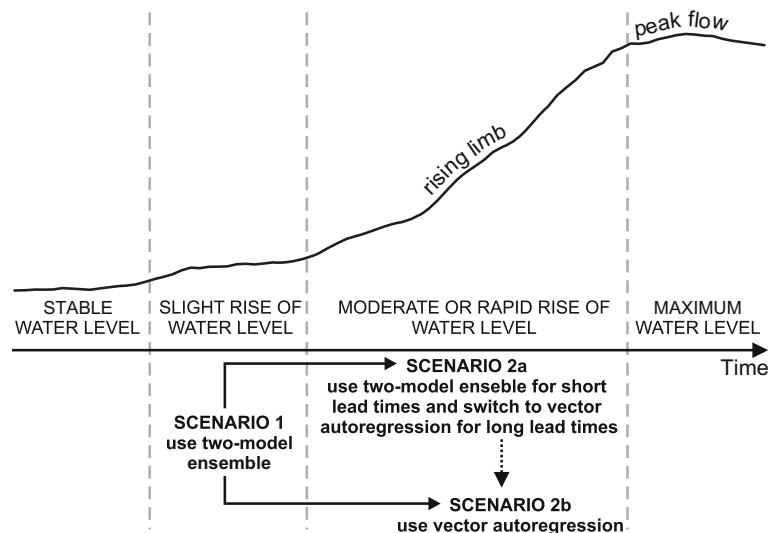
with the VAR-based forecast being slightly closer to observed data than the prediction computed using the AR method. Since weights tend to favour the AR model for short and long lead times (Fig. 9a), the resulting two-model

ensemble underestimates the hydrograph as the AR prognosis does. However, as departures from observed water level is the smallest for the ENS-bases prediction and for all lead times (Fig. 9b), and hence the scenario 1 is followed.

Table 4 Availability of ensemble members for 12 high flows

| Date of high flow episode | Ensemble members | | | | | |
|---------------------------|------------------|----|-----|----------|-------|----------|
| | VAR | AR | ACV | TOPMODEL | ANNAR | ANNARreg |
| 14/09/2013 | + | + | + | – | – | – |
| 09/12/2013 | + | + | + | – | – | – |
| 23–24/03/2014 | + | + | + | – | + | + |
| 16–17/05/2014 | + | + | + | – | + | + |
| 27–28/05/2014 | + | + | + | – | + | + |
| 31/07/2014 | + | + | + | + | + | + |
| 01–02/09/2014 | + | + | – | + | + | + |
| 11–12/09/2014 | + | + | – | + | + | + |
| 22–23/10/2014 | + | + | – | + | + | + |
| 19–21/12/2014 | + | + | + | + | + | + |
| 08–09/01/2015 | + | + | + | + | + | + |
| 29–31/03/2015 | + | + | + | – | + | + |

VAR vector autoregressive model, AR autoregressive model, ACV autocovariance method, *TOPMODEL* TOPography-based hydrological MODEL, ANNAR artificial neural network autoregressive model, ANNARreg artificial neural network autoregressive model with regularization

Fig. 5 Conceptual sketch showing the order of using scenarios 1, 2a and 2b against a background of hydrograph phases during the development of a high flow

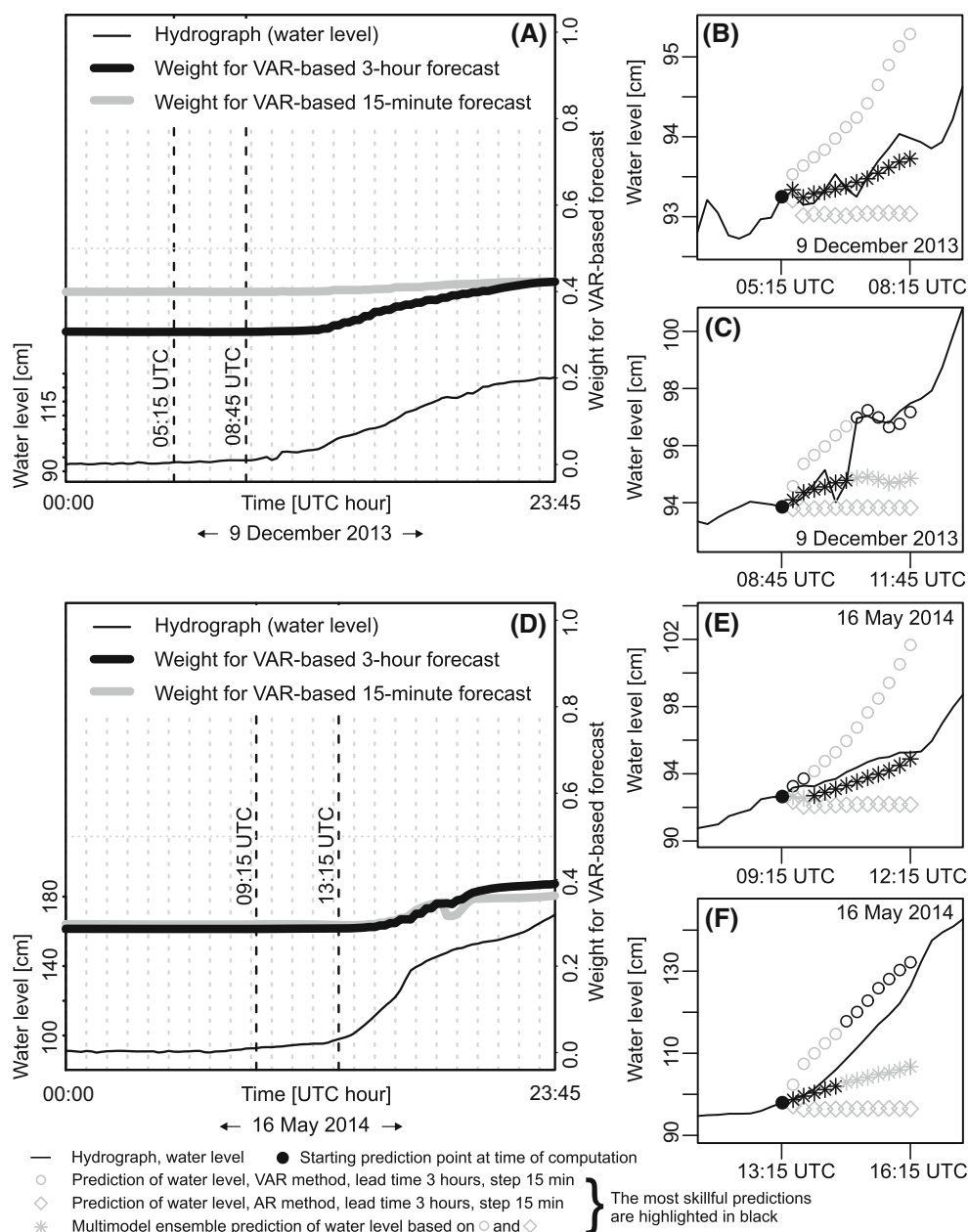
4.1.2 Scenario 1 for 50–74 % of lead times

The VAR-based prognosis issued at 01:30 UTC on 14 September 2013 revealed no mismatch for lead times of 105–180 mins, and only small overestimation for shorter lead times. The AR forecast highly underestimated the hydrograph. Due to approximately equal weights (Fig. 9d), the ENS prediction was placed in the middle of the very accurate VAR-based solution and the inaccurate AR forecast (Fig. 9e). As a consequence, the two-model ensemble was the most skillful only for lead times ranging from 15 to 90 mins. However, for lead times of 105–180 mins the ENS solution was inaccurate due to the above-mentioned perfect performance of the VAR-based prognosis for such lead times. Thus, the conditions of the scenario 1 is met

since the ENS-based prognoses are found to be most accurate for 50 % of the studied lead times (Table 5).

The VAR and AR methods applied at 14:30 UTC on 23 March 2014 led to spread of prognoses around the hydrograph, with the VAR-based forecasts departing less significantly from the hydrograph than the AR predictions. This slight supremacy of VAR over AR, when associated with almost equal weights for the two solutions (Fig. 10a), led to computation of very accurate two-model predictions for the majority (58 %) of lead times (Fig. 10b). Similarly, according to Table 5, the two-model ensemble served as the best method for predicting a gently rising water level from 20:00 UTC on 19 December 2014 for 58 % of lead times. The superior performance, with nearly perfect fit to true hydrograph, occurred for lead times ranging from 90 to 180

Fig. 6 Performance of the VAR, AR and ENS approaches before high flow (slightly rising water level) and at the initial phase of high flow (moderately or rapidly rising water level within a rising limb of hydrograph) on 9 December 2013 and 16 May 2014 at Kłodzko gauge; hydrograph and weights for VAR-based predictions with vertical dashed lines presenting two example starting prediction times (a, d), zoomed-in hydrograph to cover the first and second starting prediction times with the corresponding 3-h prognoses based on the above-mentioned methods (b, c, e, f)



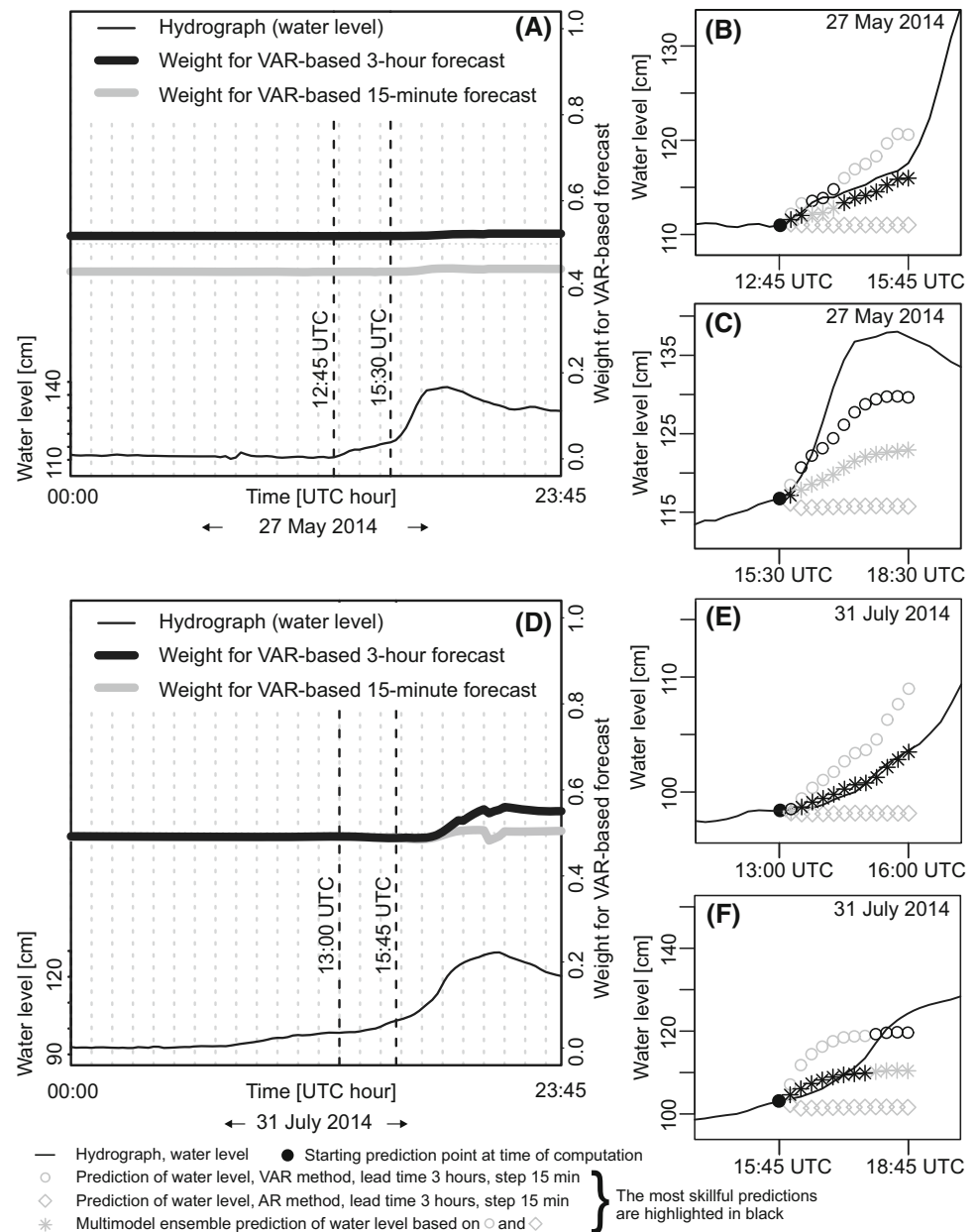
mins (Fig. 10e). Weighting procedure assigned nearly equal weights for VAR and AR approaches (Fig. 10d) which, together with similar spread of the two predictions around the true hydrograph, led to the computation of very accurate two-model ensemble prognosis for lead times ≥ 1.5 h.

A similar situation, with the two-model ensemble prognosis being the most accurate for lead times greater than 1.5 h (following Table 5, for 50 % of lead times), occurred while forecasting a slowly rising water level from 01:00 UTC on 9 January 2015 onwards (Fig. 11b). Such a performance for long lead times was due to weights that favoured the VAR-based prognosis (Fig. 11a).

4.1.3 Departures from scenario 1

In accordance with Table 5, the only case which does not fit the scenario 1 is the high flow that began 22 October 2014. It is apparent from Fig. 11e that the VAR-based prognosis issued at 22:45 UTC on 22 October 2014 perfectly fitted the true hydrograph for all lead times. The two-model ensemble underestimated the observed water level due to underestimation offered by the AR forecast. Such a setup, independently of weights, might not result in improving the VAR solution by the ENS-based prognosis

Fig. 7 Same as in Fig. 6, but for the high flows on 27 May 2014 and 31 July 2014



4.2 Prediction of rising limb of hydrograph

All studied exercises targeted at predicting a rising limb of hydrograph (12 of 12 cases) fit the second of the above-mentioned scenarios (Table 5). Within the scenario 2, seven exercises follow the scenario 2a, while the remaining five cases fit the concept of scenario 2b.

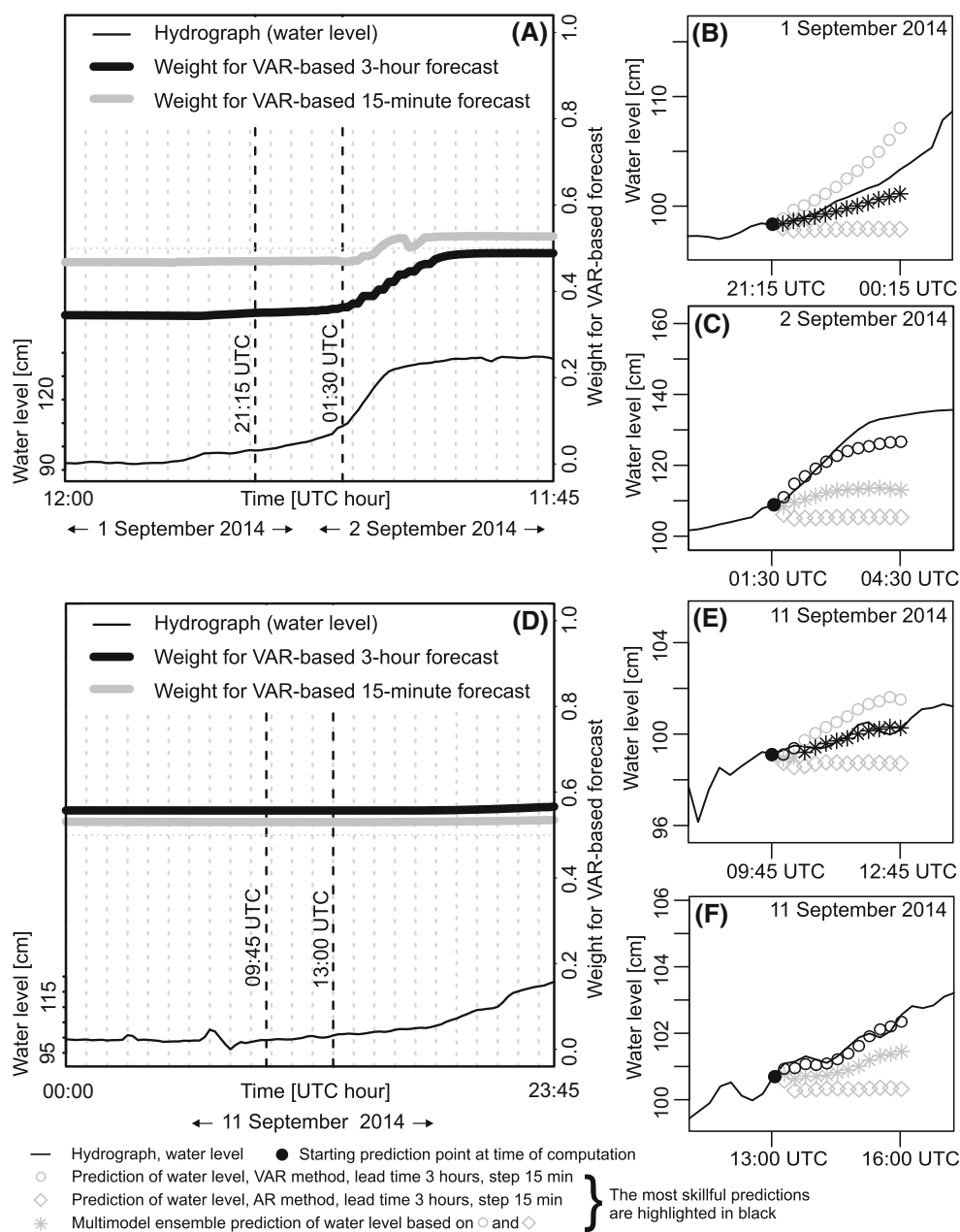
4.2.1 Scenario 2a

Figure 6c shows that in order to predict water level from 08:45 UTC on 9 December 2013 in 3-h time horizon it is also recommended to use the two-model ensemble (lead

times from 15 to 90 mins) and subsequently utilize the VAR solution (lead times from 105 to 180 mins). Although the water level rise over 3 h (08:45–11:45 UTC) was gentle (approximately 4 cm), the shape of hydrograph for lead times ≥ 105 mins was perfectly anticipated. Prognoses of a rising limb of hydrograph issued at 13:15 UTC on 16 May 2014 revealed similar characteristics (Fig. 6f). For short lead times (ranging from 15 to 75 mins) the ENS approach was found to be most skillful, while for long lead times (90–180 mins) the VAR method is recommended.

Figure 7c presents the performance of three studied prediction methods run at 15:30 UTC on 27 May 2014. The

Fig. 8 Same as in Fig. 6, but for the high flows on 1/2 September 2014 and 11 September 2014

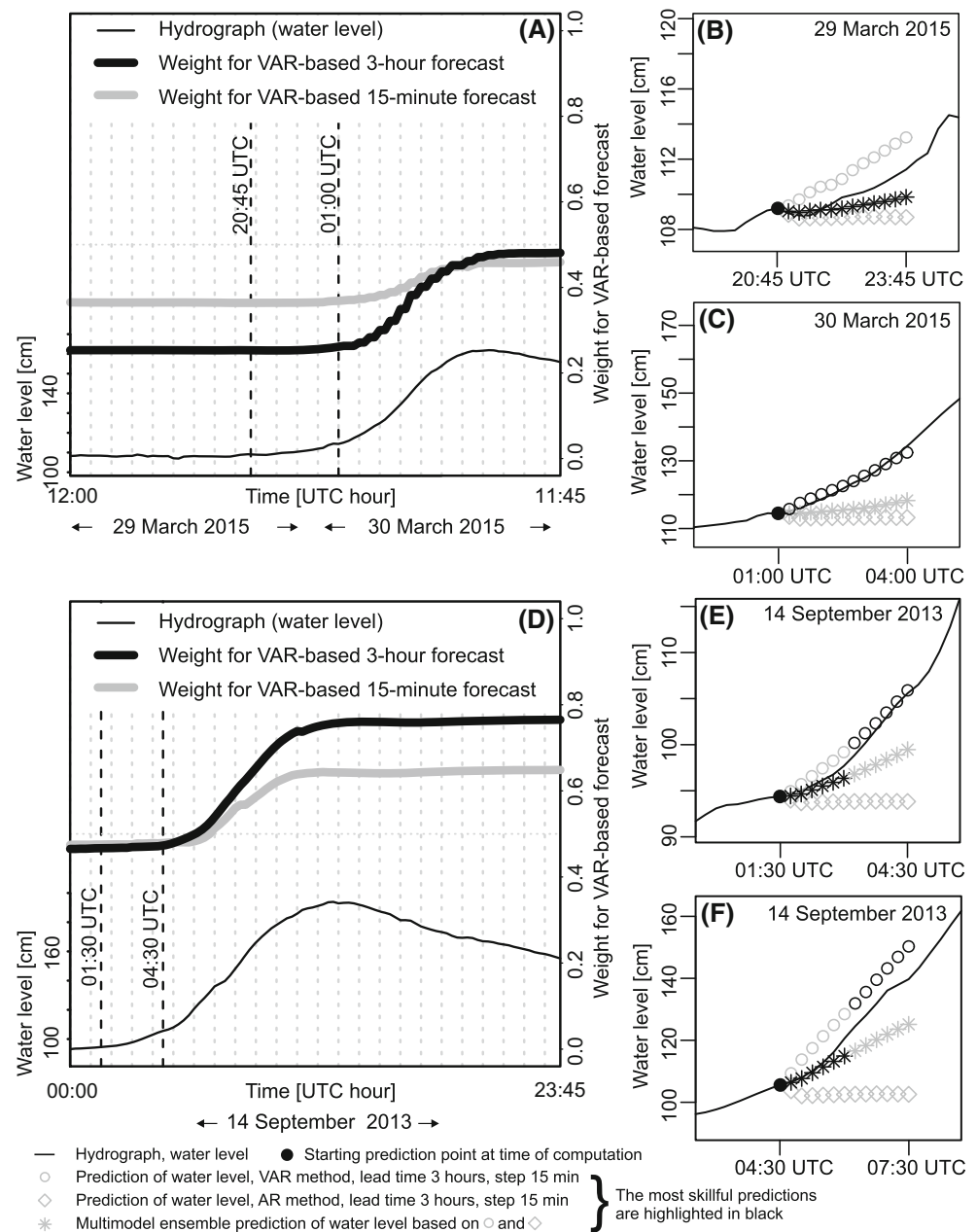


picture is rather different from those described above as the two-model ensemble was found to be the most skillful prognosis only for 15 mins into the future. For lead times from 30 to 180 mins the VAR-bases forecast was shown to perform best. Despite an underestimation of peak flow by approximately 10 cm, the VAR approach successfully imitated a rapidly rising limb of hydrograph as well as its shape.

It is apparent from Fig. 7f that the most accurate prognosis of water level issued at 15:45 UTC on 31 July 2014 should be composed of the two-model ensemble (lead times ranging from 15 to 120 mins) and the vector autoregression (lead times from 135 to 180 mins). The

latter correctly anticipated the 20-cm rise of water level in 3 h. Similarly, Fig. 9f shows that the most skillful predictions of water level issued at 04:30 UTC on 14 September 2013 were: two-model ensemble prognosis (lead times from 15 to 90 mins) and the VAR forecast (lead times from 105 to 180 mins). The VAR approach successfully predicted a rapid rise of water level (approximately 40 cm in 3 h). Likewise, the ENS-based forecast issued at 00:00 UTC on 20 December 2014 is the most skillful for very short lead times (15–30 mins), and the VAR method is recommended for lead times spanning the time horizon from 45 to 180 mins (Fig. 10f). Predictions of a rising limb of hydrograph computed at 01:15 UTC on 23 October 2014

Fig. 9 Same as in Fig. 6, but for the high flows on 29/30 March 2015 and 14 September 2013



followed the similar two-method setup (Fig. 11f). Firstly, for lead times from 15 to 45 mins the two-model ensemble works well, and subsequently, for lead times ranging from 60 to 180 mins, the VAR-based prognosis is the most skillful among the three studied methods.

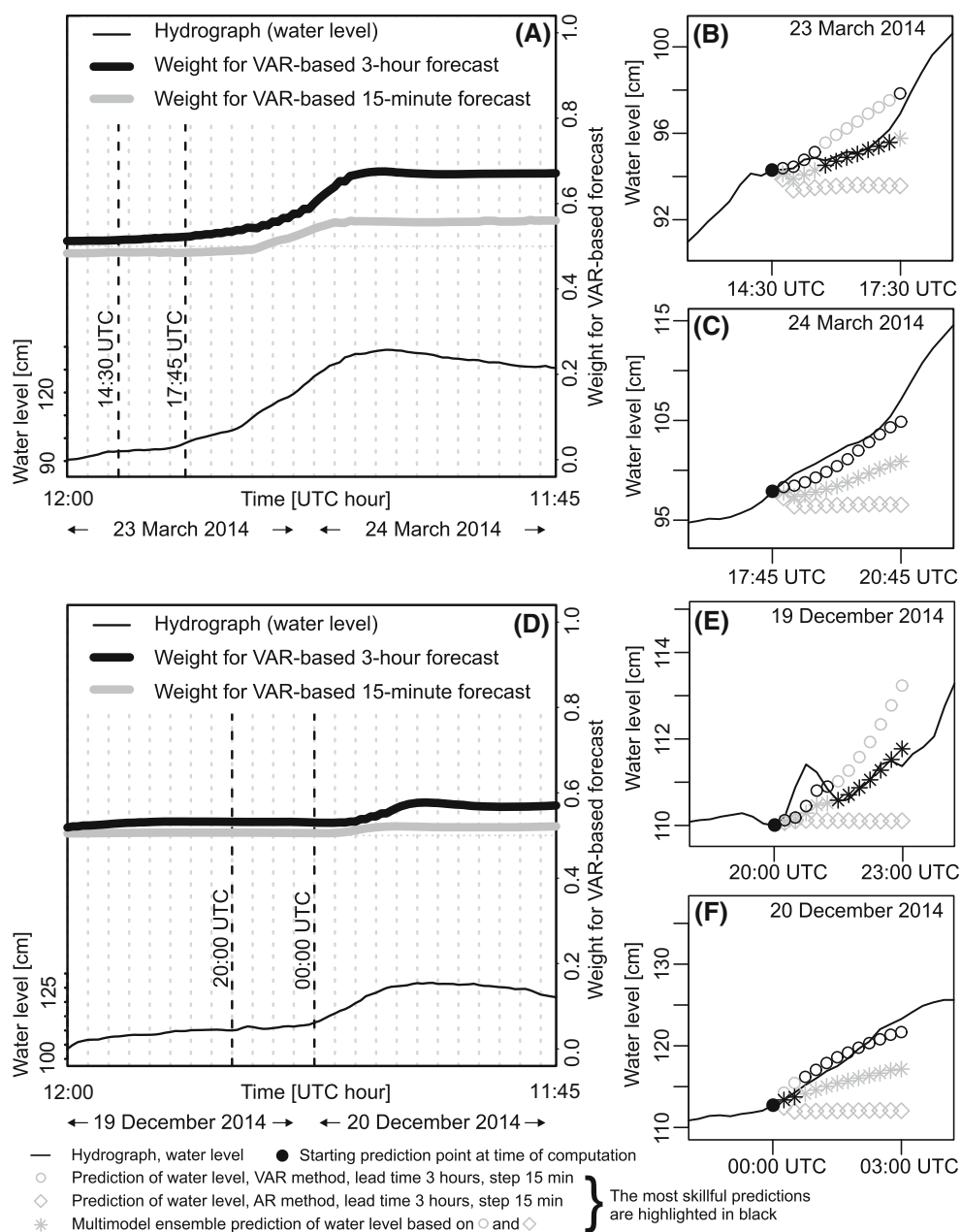
A note should be given on the chronological order of change from ENS to VAR when selecting the most skillful prediction method. In all of the studied cases, the two-model ensemble was found to be superior over the remaining methods for short, consecutive lead times. After the accuracy of the ENS-based prognosis deteriorates, the VAR solution always begins to produce the most skillful

forecast, and the latter is true for subsequent, and also consecutive lead times.

4.2.2 Scenario 2b

The VAR method itself correctly anticipated the rapid rise of water level (approximately 20 cm in 3 h) that occurred from 01:30 to 04:30 UTC on 2 September 2014 (Fig. 8c). As the VAR-based prognosis perfectly fitted or slightly overestimated the hydrograph for all lead times, the two-model ensemble was unable to improve the accuracy of the vector autoregression. Similarly, Fig. 8f shows that the

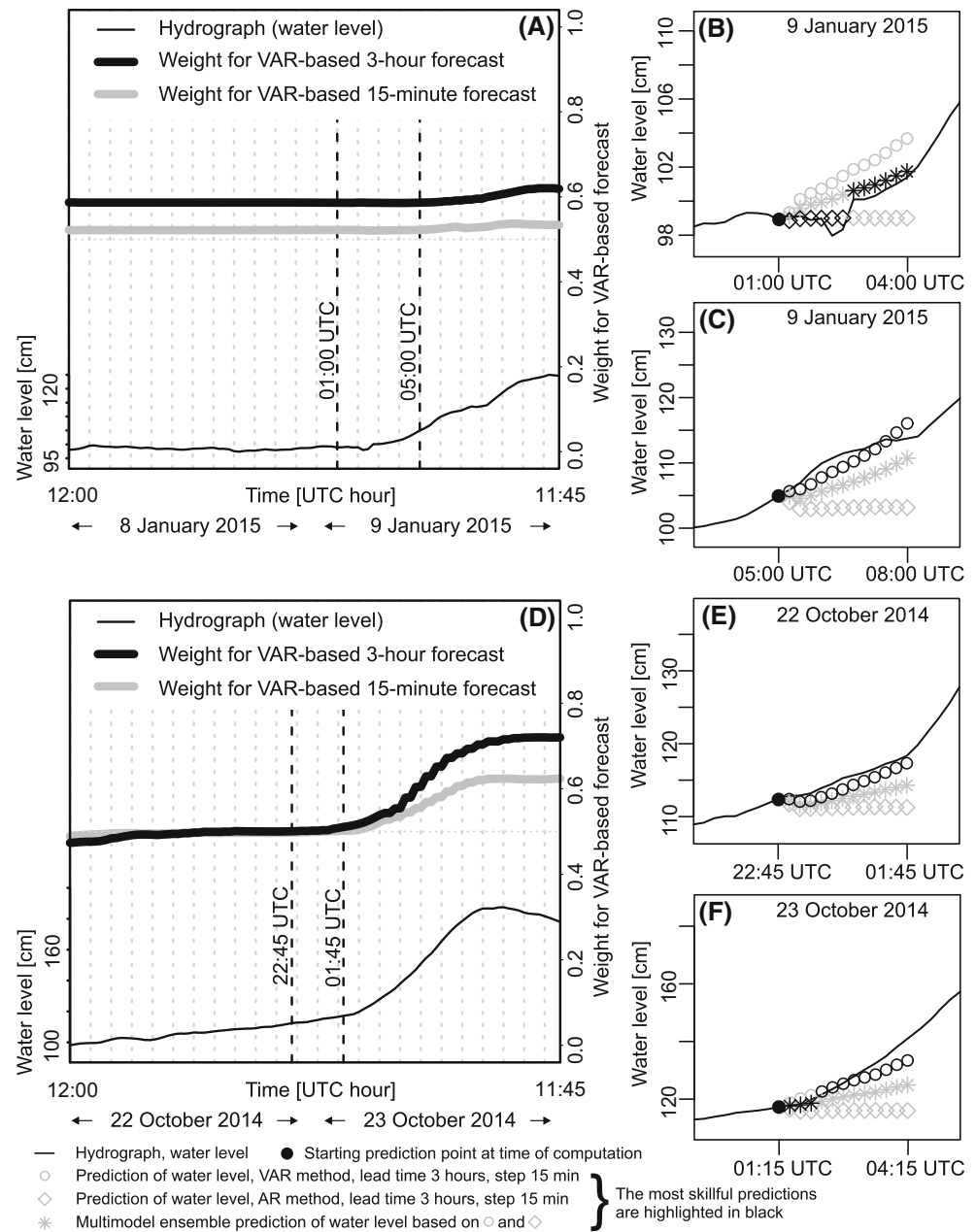
Fig. 10 Same as in Fig. 6, but for the high flows on 23 March 2014 and 19/20 December 2014



VAR method itself was also found to be the most skillful in predicting the slowly rising water level (beginning of rising limb of hydrograph) from 13:00 UTC on 11 September 2014 onwards (for all lead times). Although water level rose by 2 cm only, the VAR-bases prognosis perfectly agreed with a complex shape of the hydrograph. Noteworthy is also the performance of the VAR-based forecast itself in predicting the hydrograph from 01:00 UTC into 04:00 UTC on 30 March 2015 (Fig. 9c). The rapid rise of water level (approximately 20 cm in 3 h) was very accurately predicted, and the perfect fit was noticed for all lead times.

Both the VAR and AR predictions issued at 17:45 UTC on 23 March 2014 underestimated the observed hydrograph, with the former being very close to true data. Hence, the two-model solution was unable to be more skillful than the the VAR-based forecast. In addition, weights for the two methods were approximately equal to 0.5 (Fig. 10a) which made the ENS prognosis fall exactly between the two solutions, leading to a meaningful underestimation of riverflow. All in all, the pure VAR solution offered the most accurate prediction that successfully imitated the hydrograph over all lead times (Fig. 10c).

Fig. 11 Same as in Fig. 6, but for the high flows on 9 January 2015 and 22/23 October 2014



Very accurate was also the VAR prognosis over all lead times in the process of predicting a moderately fast rise of water level (approximately 10 cm) from 05:00 UTC to 08:00 UTC on 9 January 2015 (Fig. 11c). The two-model ensemble was unable to improve the accuracy offered by the vector autoregression since the latter model did not overestimate the hydrograph.

4.3 Experiment

Intrinsically, the question arises whether the two-model ensemble prediction switched with the VAR-based

prognoses (following the scenarios 1 and 2a/2b) is superior over multimodel ensemble predictions based on a bigger number of ensemble members. In order to address this problem, a simple experiment has been proposed. As already mentioned in Sect. 3.2, there have been six ensemble members in the HydroProg-Kłodzko prototype over the study period, however their work was discontinuous (Table 4). Following Table 4, which also contains explicit names of the ensemble members which are both data- and physically-based models, we extracted high flow events when all six models were operationally available (31 July 2014, 19–20 December 2014, 9 January 2015).

Table 5 Occurrence of scenarios 1 and 2a/2b for 12 high flows

| High flow | Before rising limb Scenario 1 | Rising limb | |
|-----------------------------|----------------------------------|-------------|-------------|
| | | Scenario 2a | Scenario 2b |
| 14 September 2013 | + (50 %) | + | – |
| 9 December 2013 | + (100 %) | + | – |
| 23/24 March 2014 | + (58 %) | – | + |
| 16 May 2014 | + (83 %) | + | – |
| 27 May 2014 | + (75 %) | + | – |
| 31 July 2014 | + (92 %) | + | – |
| 1/2 September 2014 | + (100 %) | – | + |
| 11 September 2014 | + (83 %) | – | + |
| 22/23 October 2014 | – (0 %) | + | – |
| 19/20 December 2014 | + (58 %) | + | – |
| 8/9 January 2015 | + (50 %) | – | + |
| 29/30 March 2015 | + (100 %) | – | + |
| Occurrence of scenarios 1/2 | 92 % | 100 % | |

+ (*p* %) scenario 1 is true for *p* % of lead times, +/– scenarios 2a or 2b are true/false

The experiment—limited to exactly the same starting prediction points as those from Figs. 7e/f, 10e/f and 11b/c—presents all ensemble members against a background of the two-model and six-model solutions (Fig. 12). It can be inferred from the experiment that the simple two-model setup integrated with the VAR model (our scenarios 1 and 2a/2b) may be more skillful than the six-model hydrologic ensemble prediction. Although the simplicity of the experiment and its execution for three high flows only, the exercise provides a confirmation that the two-model setup may serve skillfully in the process of forecasting initial phases of high flows.

4.4 Interpretation

The scenarios outlined above, labeled as 1 and 2a/2b, were confirmed by the above-mentioned experiments. It has been observed that there is a switch between the two-model ensemble and the vector autoregressive prognosis. Figure 5 shows a hypothetical picture that can be inferred from the case studies. It is believed that when water level starts to rise slowly, the two-model ensemble should be used (scenario 1) and—when water level begins to increase moderately or rapidly around the inflection point of a rising limb of hydrograph—two options are possible. Recommended is either the two-model ensemble for short lead times followed by the vector autoregression for long lead times (scenario 2a) or the vector autoregression itself over all lead times (scenario 2b). It is likely, but not verified in this paper, that along with the development of a high flow scenario 2a should be substituted by scenario 2b. The experiment

characterized in Sect. 4.3 and presented in Fig. 12 confirms the skillfulness of the recommended approach.

It is apparent from the analysis presented in this paper that—in the simple two-model setup—recommendation as the choice of the most appropriate predictive model should be dependent on phase of a high flow. In general, for slightly rising water level the two-model ensemble is recommended which, for a subsequently rapidly rising limb of hydrograph, should be replaced by the vector autoregressive model (Fig. 5). Although the finding seems to be robust as the scenarios repeat over different high flow events (Table 5), it is necessary to mention some limitations of the approach.

Firstly, the entire analysis was performed for one gauge, however such a choice was made on purpose to provide a superior explanatory information for vector autoregression from upstream sites (Figs. 1, 3, 4; Table 2). The selection ensured relatively long delays of the hydrological signals between gauges (Fig. 4), which is essential to build a valuable vector autoregressive model (Niedzielski 2007). It is thus believed that the site is optimally chosen to guarantee the wealth of explanatory information from the well-instrumented contributing basin (Table 2), and hence most of the model assumptions are met. Thus, due to the site-specific case study for Kłodzko it is advised not to generalize the results to dissimilar climatic zones.

Secondly, the analysis was carried out for the selected twelve high flow events. However, it is believed that the sample is big enough to carry out inference and formulate some generalizations. Noteworthy is also that fact that the selection of high flow episodes was based on a statistical

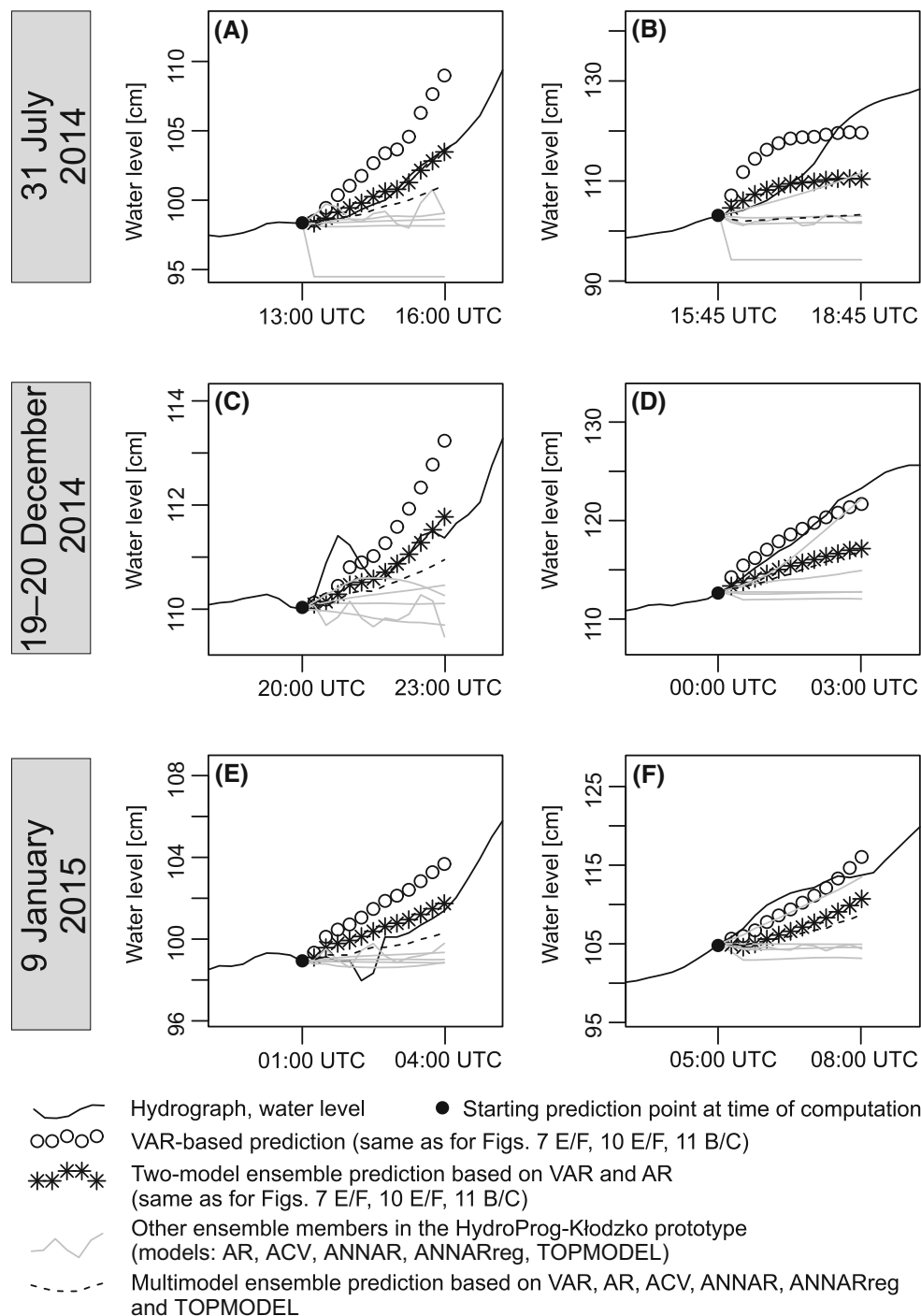


Fig. 12 Performance of the VAR-based predictions and the two-model ensemble prognoses (from scenarios 1 and 2a/2b) against a background of all ensemble members available in the HydroProg-Kłodzko prototype for three dates when all prediction models were working (see Table 4)

criterion (5 % quantile). The high flows are also presented against a background of the entire hydrograph recorded for Kłodzko over the entire 2-year study period (Fig. 13). It is apparent from this figure that key high flows have been selected, and the performance of the methods in complex hydrological situations have been presented.

A note should also be given on weights which were used to produce the two-model ensembles. At the studied starting prediction times the weights were found to be rather stable, not reacting on the forthcoming peak flows (Figs. 6a/d, 7a/d, 8a/d, 9a/d, 10a/d, 11a/d). However, it can be inferred from these figures that weights become to be

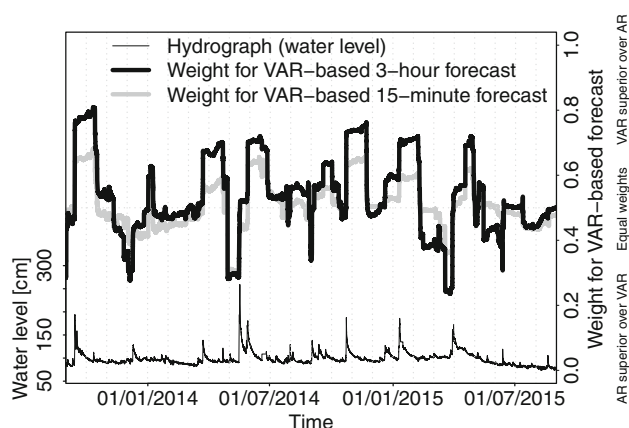


Fig. 13 Weights for the VAR-based 15-min and 3-h predictions and the hydrograph for the Kłodzko gauge in the period 01/09/2013–31/08/2015

updated around the peak hydrograph and remain modified during the recession limb of hydrograph (Fig. 13).

5 Conclusions

According to Bohn et al. (2010), there is no unequivocal agreement as to whether multimodel ensemble predictions are always more skillful than individual ensemble members. In this paper, the authors attempted to investigate this problem by examining a simple two-model setup, in which individual 3-h forecasts were based on the vector autoregressive model and the univariate autoregressive model (Niedzielski 2007). The HydroProg-Kłodzko prototype enables the real-time integration of these prognoses, with the 15-min update, leading to the real-time calculation of the two-model ensemble prediction. The system itself serves as an experimental tool for producing early warnings against high flows, and is available to the public online at <http://www.klodzko.hydroprog.uni.wroc.pl/>. Over 2 years of its experimental work, a large database of predictions and their accuracies became available which made the present analysis possible. The paper presents a detailed assessment of the prediction performance offered by the above-mentioned two data-based models and their two-model ensemble during the most significant high flows that occurred over 2 years of the HydroProg-Kłodzko experimental work.

The main finding of this paper is that, in the two-model setup outlined above, the vector autoregressive forecast or the two-model ensemble prediction are the most skillful, and their performance depends on the starting prediction times in the following way.

1. When forecasts are issued before a high flow to predict a slightly rising water level in the 3-h horizon with the

15-min temporal resolution, usually the two-model ensemble approach is very accurate for at least 50 % of lead times (this situation was referred to as scenario 1).

2. When forecasts are issued at the initial phase of a high flow to predict a rapidly rising limb of hydrograph in the 3-h horizon with the 15-min temporal resolution, usually either
 - (a) the two-model ensemble approach is accurate for short lead times (up to 1.5 h into the future) and the vector autoregressive method is accurate for longer lead times (this situation was referred to as scenario 2a) or
 - (a) the vector autoregressive model is superior over both the univariate autoregressive model and the two-model ensemble for all lead times (this situation is called scenario 2b).

Along these lines, the authors formulated a scheme which relates the order of occurrence of these scenarios to phases of high flow development. It is believed that for slightly rising water level the two-model ensemble should be recommended which, for a subsequent moderate or rapid rise of water level (rising limb of hydrograph), should be replaced by the vector autoregressive model. The switch between the two may follow the transition (1) from scenario 1 to scenario 2a (and probably to scenario 2b afterwards) or (2) directly from scenario 1 to scenario 2b. The skillfulness of the approach is confirmed in a simple experiment that compares the recommended solution with several additional ensemble members and their multimodel ensemble (Fig. 12). The results reported in this paper are site-specific and their generalization, e.g. to other climatic zones, should be carried out with caution.

The findings and scenarios confirm the results obtained earlier by Bohn et al. (2010). Indeed, the multimodel ensemble forecasts cannot be treated as the “global” most skillful prediction method, which was earlier postulated by Duan et al. (2007) and Velázquez et al. (2011) for the EPS-based hydrologic ensembles. However, the present paper shows that it is possible to identify phases of high flows in which multimodelling offers the best predictive performance and other phases in which a specific ensemble member should be used instead. The results reported in this paper are preliminary, and hence further investigation is needed. A natural continuation may be a confirmation of the transition from scenario 2a to scenario 2b as well as estimation of the moment when it is recommended to switch from scenario 1 to scenario 2a/2b.

Acknowledgments The research has been financed by the National Science Centre, Poland, through the Grant No. 2011/01/D/ST10/04171 under leadership of Dr hab. Tomasz Niedzielski, Professor at the University of Wrocław, Poland. The authors kindly acknowledge

the authorities of the County Office in Kłodzko for productive partnership and making the data of the Local Flood Monitoring System (Lokalny System Oslony Przeciwpowodziowej—LSOP) available for scientific purposes. The authors are grateful to Dr Danuta Trojan for discussions on LSOP and on water management problems in Kłodzko County. We also thank Dr hab. Mariusz Szymanowski for preparing the Digital Elevation Model that has been used herein for production of the location map. We are also indebted to Dr Małgorzata Wiczorek and Dr Waldemar Spallek who kindly provided the additional spatial data which have also been utilized in the process of production of the above-mentioned map. We also thank Dr hab. Maciej Kryza, Professor at the University of Wrocław, and Dr Paweł Netzel who contributed to the design of the HydroProg system. Prognoses of rainfall, operationally computed using the infrastructure of Wrocław Centre for Networking and Supercomputing (Wrocław University of Technology), have been kindly provided by Dr hab. Maciej Kryza, Professor at the University of Wrocław. The detention reservoir data have been offered by dam managers. The engine of the HydroProg-Kłodzko prototype is based on the following open source solutions: R, PostgreSQL and MySQL. Real-time riverflow data are available courtesy of the County Office in Kłodzko, Poland.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Ajami NK, Duan Q, Gao X, Sorooshian S (2006) Multi-model combination techniques for hydrological forecasting: application to distributed model intercomparison project results. *J Hydrometeorol* 7:755–768
- Akaike H (1971) Autoregressive model fitting for control. *Ann Inst Stat Math* 23:163–180
- Alfieri L, Pappenberger F, Wetterhall F, Haiden T, Richardson D, Salamon P (2014) Evaluation of ensemble streamflow predictions in Europe. *J Hydrol* 517:913–922
- Bohn TJ, Sonessa MY, Mergia Y, Lettenmaier DP (2010) Seasonal hydrologic forecasting: Do multimodel ensemble averages always yield improvements in forecast skill? *J Hydrometeorol* 11:1358–1372
- Brockwell PJ, Davis RA (1996) Introduction to time series and forecasting. Springer, New York
- Buizza R, Houtekamer PL, Toth Z, Pellerin G, Wei M, Zhu Y (2005) A comparison of the ECMWF, MSC, and NCEP global ensemble prediction systems. *Mon Weather Rev* 133:1076–1097
- Butts MB, Overgaard J, Dubicki A, Strońska K, Lewandowski A, Olszewski T, Kolerski T (2006) Intercomparison of distributed hydrological models for flood forecasting in the Odra River basin. *Geophys Res Abstr* 8:07250
- Butts M, Dubicki A, Strońska K, Jrgensen G, Nalberczynski A, Lewandowski A, Van Kalken T (2007) Flood forecasting for the upper and middle Odra River basin. In: Begum S et al (eds) *Flood risk management in Europe*. Springer, Dordrecht, pp 353–384
- Cloke HL, Pappenberger F (2009) Ensemble flood forecasting: a review. *J Hydrol* 375:613–626
- Cloke HL, Pappenberger F, van Andel SJ, Schaake J, Thielen J, Ramos M-H (2013) Hydrological ensemble prediction systems. *Hydrol Process* 27:1–4
- Dechant C, Moradkhani H (2011a) Radiance data assimilation for operational snow and streamflow forecasting. *Adv Water Resour* 34:351–364
- DeChant CM, Moradkhani H (2011b) Improving the characterization of initial condition for ensemble streamflow prediction using data assimilation. *Hydrol Earth Syst Sci* 15:3399–3410
- DeChant CM, Moradkhani H (2015) Analyzing the sensitivity of drought recovery forecasts to land surface initial conditions. *J Hydrol* 526:89–100
- Demargne J, Wu L, Regonda SK, Brown JB, Lee H, He M, Seo D-J, Hartman R, Herr HD, Fresch M, Schaake J, Zhu Y (2013) The science of NOAA's operational hydrologic ensemble forecast service. *Bull Am Meteorol Soc* 95:79–98
- Duan Q, Ajami NK, Gao X, Sorooshian S (2007) Multi-model ensemble hydrologic prediction using Bayesian model averaging. *Adv Water Resour* 30:1371–1386
- Dubicki A, Malinowska-Malek J, Strońska K (2005) Flood hazards in the upper and middle Odra River basin—a short review over the last century. *Limnologica* 35:123–131
- Franz K, Ajami N, Schaake J, Buizza R (2005) Hydrologic ensemble prediction experiment focuses on reliable forecasts. *Eos* 86:239
- Gericke OJ, Smithers JC (2014) Review of methods used to estimate catchment response time for the purpose of peak discharge estimation. *Hydrol Sci J* 59:1935–1971
- Kasprzak M (2010) Wezbrania i powodzie na rzekach Dolnego Śląska. In: Migoń P (ed) *Wyjątkowe zdarzenia przyrodnicze na Dolnym Śląsku i ich skutki*, *Rozprawy Naukowe Instytutu Geografii i Rozwoju Regionalnego Uniwersytetu Wrocławskiego* 14, pp 81–140
- Kumar A, Singh R, Pratyasha Jena P, Chatterjee C, Mishra A (2015) Identification of the best multi-model combination for simulating river discharge. *J Hydrol* 525:313–325
- Lardies J (1996) Analysis of a multivariate autoregressive process. *Mech Syst Signal Process* 10:747–761
- Leedal D, Weerts AH, Smith PJ, Beven KJ (2013) Application of data-based mechanistic modelling for flood forecasting at multiple locations in the Eden catchment in the National Flood Forecasting System (England and Wales). *Hydrol Earth Syst Sci* 17:177–185
- Lenar-Matyas A, Poulard C, Ratomski J, Royet R (2009) The construction and action of dry dams with the different characteristics and location. *Infrastruktura i Ekologia Terenów Wiejskich (Infrastructure and Ecology of Rural Areas)* 9:115–129
- Li H, Luo L, Wood EF, Schaake J (2009) The role of initial conditions and forcing uncertainties in seasonal hydrologic forecasting. *J Geophys Res* 114:D04114. doi:10.1029/2008JD010969
- McCuen RH (2003) Modeling hydrologic change—statistical methods. Lewis Publishers, Boca Raton
- Najafi MR, Moradkhani H (2015) Multi-model ensemble analysis of runoff extremes for climate change impact assessments. *J Hydrol* 525:352–361
- Neumaier A, Schneider T (2001) Estimation of parameters and eigenmodes of multivariate autoregressive models. *ACM Trans Math Softw* 27:27–57
- Niedzielski T (2007) A data-based regional scale autoregressive rainfall-runoff model: a study from the Odra River. *Stoch Environ Res Risk Assess* 21:649–664
- Niedzielski T (2010) Empirical hydrologic predictions for southwestern Poland and their relation to ENSO teleconnections. *Artif Satell* 45:11–26
- Niedzielski T, Miziski B, Kryza M, Netzel P, Wiczorek M, Kasprzak M, Kosek W, Migo P, Szymanowski M, Jeziorska J, Witek M (2014) HydroProg: a system for hydrologic forecasting in real time based on the multimodelling approach. *Meteorol Hydrol Water Manag Res Oper Appl* 2:65–72

- Olsson J, Lindström G (2008) Evaluation and calibration of operational hydrological ensemble forecasts in Sweden. *J Hydrol* 350:14–24
- Ramos H-H, Bartholmes J, Thielen-del Pozo J (2007) Development of decision support products based on ensemble forecasts in the European flood alert system. *Atmos Sci Lett* 8:113–119
- Raynaud D, Thielen J, Salamon P, Burek P, Anquetin S, Alfieri L (2015) A dynamic runoff co-efficient to improve flash flood early warning in Europe: evaluation on the 2013 central European floods in Germany. *Meteorol Appl* 22:410–418
- Regonda SK, Seo D-J, Lawrence B, Brown JD, Demargne J (2013) Short-term ensemble streamflow forecasting using operationally-produced single-valued streamflow forecasts—a hydrologic model output statistics (HMOS) approach. *J Hydrol* 497:80–96
- Rosner B (1983) Percentage points for a generalized ESD many-outlier procedure. *Technometrics* 25:165–172
- Schaake J, Franz K, Bradley A, Buizza R (2006) The hydrologic ensemble prediction experiment (HEPEX). *Hydrol Earth Syst Sci Discuss* 3:3321–3332
- Schaake J, Pailleux J, Thielen J, Arritt R, Hamill T, Luo L, Martin E, McCollor D, Pappenberger F (2010) Summary of recommendations of the first workshop on postprocessing and downscaling atmospheric forecasts for hydrologic applications held at Mto-France, Toulouse, France, 1518 June 2009. *Atmos Sci Lett* 11:59–63
- See L, Abrahart RJ (2001) Multi-model data fusion for hydrological forecasting. *Comput Geosci* 27:987–994
- Seiller G, Anctil F, Perrin C, Gelfan A (2012) Multimodel evaluation of twenty lumped hydrological models under contrasted climate conditions. *Hydrol Earth Syst Sci* 16:1171–1189
- Sen AK, Niedzielski T (2010) Statistical characteristics of riverflow variability in the Odra River basin, Southwestern Poland. *Pol J Environ Stud* 19:387–397
- Smith PJ, Beven KJ, Leedal D, Weerts AH, Young PC (2014) Testing probabilistic adaptive real-time flood forecasting models. *J Flood Risk Manag* 7:265–279
- Thielen J, Schaake J, Hartman R, Buizza R (2008) Aims, challenges and progress of the hydrological ensemble prediction experiment (HEPEX) following the third HEPEX workshop held in Stresa 27 to 29 June 2007. *Atmos Sci Lett* 9:29–35
- van Andel SJ, Weerts A, Schaake J, Bogner K (2013) Post-processing hydrological ensemble predictions intercomparison experiment. *Hydrol Process* 27:158–161
- Velázquez JA, Anctil F, Ramos MH, Perrin C, Burek P (2011) Can a multi-model approach improve hydrological ensemble forecasting? A study on 29 French catchments using 16 hydrological model structures. *Adv Geosci* 29:33–42
- Werner M, Schellekens J, Gijsbers P, van Dijk M, van den Akker O, Heynert K (2013) The Delft-FEWS Cōw forecasting system. *Environ Model Softw* 40:65–77
- Wood AW, Lettenmaier DP (2008) An ensemble approach for attribution of hydrologic prediction uncertainty. *Geophys Res Lett* 35:L14401. doi:[10.1029/2008GL034648](https://doi.org/10.1029/2008GL034648)
- Yan H, Moradkhani H (2016) Toward more robust extreme flood prediction by Bayesian hierarchical and multimodeling. *Nat Hazards* 81:203–225
- Zappa M, Rotach MW, Arpagaus M, Dorninger M, Hegg C, Montani A, Ranzi R, Ament F, Germann U, Grossi G, Jaun S, Rossa A, Vogt S, Walser A, Wehrhan J, Wunram C (2008) MAP D-PHASE: real-time demonstration of hydrological ensemble prediction systems. *Atmos Sci Lett* 9:80–87